

REMARKS

The following claim amendments are requested:

Claims Pending before Preliminary Amendment: 1-74

Please Cancel Claims Without Prejudice:

3, 6, 10, 12-15, 17-20, 24-27, 29-50, 54-70

Claims Amended: 1, 2, 9, 11, 21, 23, 28

Claims Added: None.

Pending claims: 1, 2, 4, 5, 7-9, 11, 16, 21-23, 28, 51-53, 71-74

An amendment consisting of a substitute specification has been provided in both clean copy and marked-up copy to satisfy a missing parts notice of September 28, 2004. Typos, where evident have been corrected. The substituted specification contains no new matter and directs a sequence listing, required under 37 CFR 1.821-1.825, into the specification.

The official sequence listing is provided on paper as an appendix to the application. The sequence listing is also provided in computer readable form (CRF) on two CDs. A statement is provided that the content of the sequence listing information recorded in CRF is identical to the written sequence listing and contains no new matter.

The claims remain unaltered and are provided only for reference, so that the right specification gets matched to the right claims.

Substitute drawings are also provided to conform to 37 CFR 1.821-1.825 and to fulfill the same missing parts request.

CONCLUSION

Should any matter in this case remain unresolved, the undersigned attorney respectfully requests a telephone conference with the Examiner to resolve any such outstanding matter.

Respectfully Submitted,

Date: 12-28-04

By: Mark C. Farrell
Lee & Hayes PLLC
Mark C. Farrell
Reg. No. 45,988
(509) 324-9256

AMENDMENTS TO THE DRAWINGS

Substitute drawings are provided in order to meet the missing parts requirements to meet 37 CFR 1.821-1.825.

Substitute drawings are also provided to conform to 37 CFR 1.821-1.825 thus fulfilling the same missing parts request. In the drawings, small letters for nucleic acid bases are substituted for the non-conforming capital letters used in the drawings being substituted for.



EV554746222

1

METHODS OF DESIGNING, SYNTHESIZING, AND PROPAGATING
REFERENCE NUCLEIC ACIDS

5

SEQUENCE LISTING

A sequence listing in accordance with 37 CFR 1.822 and WIPO Standard ST.25 and produced by PATENTIN software version 3.3 is provided beginning on a separate sheet, on four sheets of paper and incorporated by reference into the specification. A sequence listing identical to the paper version is provided in duplicate in computer readable form (CRF) on two CDs, each containing a single file "SH1-0001US.ST25.TXT" of size 3.87 kilobytes that is incorporated by reference into the specification and is an exact duplicate of the sequence listing provided on paper. The file SH1-0001US.ST25.TXT is compatible with IBM-PC, MS-DOS, and MS-Windows applications.

TECHNICAL FIELD

The subject matter relates generally to molecular biology and more specifically to methods of designing, synthesizing and propagating reference nucleic acids.

BACKGROUND

Medical diagnostic tests in clinical laboratories commonly require stringent quality control as mandated by government agencies and standards organizations. The National Committee for Clinical Laboratory Standards (NCCLS) suggests accreditation guidelines that include calibrating equipment against control samples and performing tests of patient samples in

tandem with consistent references (NCCLS, Villanova, Pennsylvania). Other organizations, such as the Joint Commission on Accreditation of Healthcare Organizations (JCAHO) and the American Society for Clinical Pathology (ASCP) also recommend or mandate standardized clinical procedures often
5 requiring updated (non-expired) and well-inventoried supplies of clinical reference reagents and controls (JCAHO, Washington D.C.; ASCP, Chicago, Illinois). Control references must be tested in conjunction with each test of a patient sample according to the Clinical Laboratory Improvement Act of 1988, which applies to over 175,000 laboratory entities (CLIA '88 is
10 described at 42 C.F.R., parts 493.1 - 493.1850). The College of American Pathologists (CAP) and the American College of Medical Genetics (ACMG) also mandate comparison with references during each patient test (CAP, Northfield, Illinois; ACMG, Bethesda, Maryland).

Clinical assays often involve either deoxyribonucleic acid (DNA) or
15 ribonucleic acid (RNA). For example, nucleic acid diagnostics may be performed to find infectious DNA or RNA from an invading virus or bacteria. Reference nucleic acids are typically used as chromatographic, spectroscopic, and/or spectrophotometric controls, for example during gel electrophoresis monitored by a laser "electric eye."

20 DNA provides a template for generating bodily proteins using sequences of four DNA bases (adenine, thymine, guanine, and cytosine). The mutation of any one base in a single-stranded DNA sequence may be enough to form a defective allele resulting in disease. Many disease moieties thus have an underlying genetic etiology.

25 A gene is a length of DNA on a chromosome associated with some particular process or characteristic of an individual. A gene is conventionally

considered a fundamental building block of heredity that determines observable characteristics, i.e., the “phenotype” of the individual organism. The underlying “chemical” genetic constitution of the individual is instead called its “genotype.”

5 Genes are observed to be lined up on human chromosomes in a sequential order. The sequential order of genes is the same for both members of a chromosome pair. Therefore genes occur in pairs (homologous genes). The two genes in a pair may occur in different forms called “alleles” and the phenotypic expression of one allele or the other in a pair depends on
10 the types of alleles present. Mutations are the changes in the DNA sequence that may convert one allele to another. An individual who carries two of the same alleles is homozygous for that gene while an individual who has two different alleles for a gene pair is heterozygous for that gene. The occurrence of mutations that are deleterious to the normal expression of an
15 allele may result in malfunction of that allele. The co-occurrence of a “normal” allele with a “mutated” (or abnormal) allele at the same gene in a heterozygous individual may result in a new (or disease) phenotype. In such a case, the mutated allele is described as acting in a dominant fashion over the normal allele. If the mutated allele does not cause any change in
20 phenotype of the heterozygous individual, but causes a change (or disease) only when the individual is homozygous for that mutated allele, the mutated allele is described as acting in a recessive fashion compared to the normal allele. Thus, dominance and recessiveness describe the relative effect of gene expression of an allele when two distinct alleles occur together.

25 Carriers of genetic diseases typically carry a heterozygous recessive allele that includes a mutation capable of causing the disease. However, the

mutated recessive allele may not be expressed in the carrier because its deleterious effect on the phenotype is masked by the normal (non-mutated) allele. Thus, a carrier may possess mutations in his genotype that can be passed down to descendants to cause the disease yet the carrier presents a normal phenotype (expressed characteristics) and is thus disease free. On the other hand, a person who experiences a genetically mediated disease may be a heterozygous “carrier” who has a mutated dominant allele for the disease. Still further, a person who experiences a disease may be a homozygous “carrier” with identical homologous genes that each has a mutation at a particular locus that causes the disease.

Given the number of different kinds of genetic diseases, the different possibilities for homozygous and heterozygous causation, and the need for both disease and carrier testing, maintaining recommended or mandated clinical supplies of high quality nucleic acid references and controls (hereinafter referred to as “reference nucleic acids” or just “references”) presents daunting challenges to genetics reference facilities and molecular diagnostic laboratories. As shown in **Fig. 1**, a reference nucleic acid 100 to be amplified for use as a test control comprises single-stranded or double-stranded reference RNA or DNA of known quantity and known quality within currently accepted tolerances. The ideal reference nucleic acid 100 to be amplified should resemble the patient sample to be tested as closely as possible and moreover, should be usable in all configurations of a given type of test. However, a reference nucleic acid 100 for clinical use may not be easily available in an adequate quantity and quality. Further, once an adequate quantity and quality of the reference nucleic acid 100 is obtained, the reference also needs to be reasonably easy to manufacture and store.

These various requirements are difficult to meet because a single patient test often includes many diverse steps, such as polymerase chain reactions, enzymatic manipulations, sequencing reactions, hybridizations, electrophoreses, etc., each placing a different demand on the reference.

- 5 Limited sources for obtaining a reference nucleic acid 100 to be amplified exacerbate a quality problem by causing a short supply leading to an increased likelihood that references of lower quality will be allowed in order to bolster the supply.

Typically, a reference nucleic acid 100 to be amplified originates from
10 a human source 102, but if not available in sufficient quantity or not amenable to storage, then conventional chemical synthesis 104 may augment or replace the human source 102. Depending on the identity of the reference nucleic acid 100 to be amplified and its origin, various methods may be needed to refine, develop and increase the supply, each method yielding a
15 product that may or may not have consistent quality with products yielded by other methods. A first conventional method 106 may merely isolate and purify the reference nucleic acid 100 to be amplified from a human source 102 and/or from conventional chemical synthesis 104. A second conventional method 108 may replicate the reference nucleic acid 100 to be
20 amplified by cloning in a vector (e.g., a plasmid) and allowing a species of bacteria to propagate the vector. A third conventional method 110 may undertake amplification of a human-derived reference nucleic acid 100 in an automated cycler. Other conventional methods not reviewed here are represented by an "Nth" conventional method 112 that yields a variable,
25 heterogeneous product. Most of these known methods yield products that may have inconsistent quality and/or stability, and many of the example

methods are cumbersome and expensive, as well as dependent on starting materials from a human source 102.

SUMMARY

Subject matter includes design, synthesis, and propagation of synthetic and semi-synthetic reference nucleic acids and mixtures of reference nucleic acids for use in tests, such as molecular screening, genetic testing, carrier testing, and
5 diagnostic assays. In one implementation, methods are described for design, synthesis, and propagation of reference nucleic acid mixtures and a system is presented for using the mixtures.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a graphic representation of conventional methods of obtaining nucleic acids.

Fig. 2 is a graphic representation of an exemplary method of producing a reference nucleic acid.

Fig. 3 is a graphic representation of an exemplary method of producing a mixture of reference nucleic acids.

Fig. 4 is a graphic representation of an ~~exemplary tagged~~ example synthetic ~~tagged~~ reference nucleic acid of SEQ ID NO: 1 tagged with example first tag SEQ ID NO: 2 and example second tag SEQ ID NO: 3, and example primer set consisting of example first primer SEQ ID NO: 4 and example second primer SEQ ID NO: 5.

Fig. 5 is a graphic representation of a polymerase chain reaction cycle using an ~~exemplary tagged~~ example synthetic ~~tagged~~ reference nucleic acid SEQ ID NO: 1 tagged with example first tag SEQ ID NO: 2 and example second tag SEQ ID NO: 3, wherein an example first primer SEQ ID NO: 4 makes complementary strand SEQ ID NO: 6 and example second primer SEQ ID NO: 5 makes copy strand SEQ ID NO: 1 plus SEQ ID NO: 2.

Fig. 6 is a graphic representation of an exemplary synthetic tagged duplex of a reference nucleic acid, synthesized as an example first sequence SEQ ID NO: 7 with an example first tag SEQ ID NO: 8 and an example second sequence SEQ ID NO: 9 with an example second tag SEQ ID NO: 2.

Fig. 7 is a graphic representation of production of an exemplary mixture of reference nucleic acids.

Fig. 8 is a graphic representation of an exemplary mixture of multiple mixtures, wherein the propagation of each of the multiple mixtures is controlled by an associated primer set.

Fig. 9 is a graphic representation of exemplary base sequence intervals providing isolation of reference mutation sequences from example
5 tags SEQ ID NO: 10 and SEQ ID NO: 11, thereby facilitating production of
multiple reference nucleic acids representing genetic mutations.

Fig. 10 is a graphic representation of exemplary base intervals providing isolation of multiple reference mutation sequences from each other
10 and from example tags SEQ ID NO: 10 and SEQ ID NO: 11, thereby
facilitating production of an exemplary reference nucleic acid representing
multiple genetic mutations.

Fig. 11 is a graphic representation of an exemplary method of ligation extension of example tagged reference nucleic acid oligonucleotides SEQ ID
15 NO: 12 and SEQ ID NO: 13 using example bridge sequence SEQ ID NO: 14
to make example reference nucleic acid sequence SEQ ID NO: 15.

Fig. 12 is a graphic representation of an exemplary method of overlap extension of example reference nucleic acid oligonucleotides SEQ ID NO:
12 and SEQ ID NO: 13 by synthesizing an example region of
20 complementation SEQ ID NO: 16 in order to make example longer
polynucleotides SEQ ID NO: 17 and SEQ ID NO: 18.

Fig. 13 is a graphic representation of exemplary ligation methods for multiple segments of reference nucleic acid.

Fig. 14 is a graphic representation of exemplary methods for making
25 synthetic and semi-synthetic homologous gene segments and alleles for
carrier testing.

Fig. 15 is a block diagram of an exemplary universal genetic mutation testing system.

DETAILED DESCRIPTION

Overview

The subject matter describes methods of designing, synthesizing and/or propagating synthetic nucleic acid sequences for use as references in molecular diagnostic and genetic testing—especially mutation testing—of human or nonhuman nucleic acids. In one aspect of the subject matter, relatively short segments of much larger genomic nucleic acids are designed, synthesized, and amplified because they have relevance as clinical references or controls—thus eliminating conventional problems associated with conventionally producing many irrelevant segments of large nucleic acids in quantity along with the few relevant segments. Because the segments that have relevance as clinical references are often relatively short, the subject matter includes synthesis of reference nucleic acids and mixtures thereof that are relatively easy to design and synthesize and capable of being manipulated to advantage during amplification, other enzymatic modification, and detection.

In one implementation, for example, the subject matter can be used to amplify an entire mixture of reference nucleic acids in concert. In another or the same implementation, the subject matter can be used to build a large, clinically relevant, nucleic acid polynucleotide in easy-to-synthesize smaller segments. These implementations, singly or in combination, allow the production of nucleic acids and nucleic acid mixtures that may be otherwise unobtainable or may be obtainable at a lower quality and at greater effort and expense by conventional methods.

Amplifying an entire mixture of reference nucleic acids in concert has multiple uses. For example, in many genetic diseases, any one of several

hundred mutations in a single gene may cause the disease. The several hundred mutations do not occur together in a real gene that exists in nature. The subject matter presented here, however, can create a single nucleic acid mixture that models, for purposes of mutation testing, all of multiple mutations as if a hypothetical gene or DNA segment having all the multiple mutations were present.

An exemplary synthesis of a reference nucleic acid mixture to be amplified (“amplified” and “propagated” are used somewhat interchangeably herein) first includes synthesizing the starting constituents to be used as “seed templates” for producing a desired reference nucleic acid mixture. In one implementation, the constituents may include a collection of polymeric or oligomeric nucleic acid sequences of interest, for example, a collection of oligonucleotides each representing a mutated section of a gene or of any DNA segment. In another or the same implementation, the constituents of the reference nucleic acid mixture are segments of a larger polynucleotide to be assembled into the larger polynucleotide by exemplary methods presented below with respect to Figures ~~9 and 10~~ 11 and 12. The subject matter that will be described below allows the designing practitioner to produce relatively complex mixtures and/or relatively complex molecules while only having to perform multiple syntheses of relatively short and simple sequences.

The production of a complex reference nucleic acid mixture and/or the assemblage of one or more large molecules can be executed via an exemplary system of “tags:” nucleic acid sequences for priming, bridging, and/or overlapping the constituent reference nucleic acids. For example, during the synthesis of each constituent reference nucleic acid a first nucleic

acid tag is bound to (e.g., synthesized onto) a first end of the particular reference nucleic acid. The first tag provides a priming site for the first of two primers for an amplification procedure. The synthesis further includes binding (e.g., synthesizing) a second tag to a second end of the reference nucleic acid to be amplified. The second tag provides a template for the first primer to create a priming site for the second primer on a complementary nucleic acid strand produced by the first primer. In other words, the first tag has a base sequence that complements a base sequence of the first primer while the second tag has the same base sequence as the second primer. In each of a series of polymerase chain reaction (PCR) amplifications, the first primer is extended into single-stranded complements of the original tagged reference nucleic acid (100 in Fig. 4) while the second primer is extended into copies of the original tagged reference nucleic acid using the single-stranded complements produced by the first primer as templates.

Thus, an exemplary method includes synthesizing multiple single-stranded reference nucleic acids and tagging each of the multiple reference nucleic acids with a copy of the above-described first nucleic acid tag and a copy of the second nucleic acid tag, i.e., on opposite and appropriate ends of each of the multiple reference nucleic acids. When the mixture of reference nucleic acids is exposed to the first and second primers in one or more PCR amplifications, the entire mixture is amplified using only the two primers. Because in this example each of the multiple reference nucleic acids, including tags, can be “human designed” and because propagation of the entire mixture is by one primer set, a high degree of process control is available to an architect of reference nucleic acids and a high degree of

quality and reproducibility is afforded to the resulting reference nucleic acid mixture.

As mentioned above, amplifying an exemplary reference nucleic acid mixture is useful in many kinds of tests. For example, in mutation testing requiring a different nucleic acid control reagent for testing each of multiple gene mutations that could be responsible for a disease, exemplary subject matter allows the multiple control reagents to be combined into one mixture and amplified indefinitely into a supply that has consistent quality and consistent proportion of constituent reference nucleic acids.

Exemplary Methods

Fig. 2 shows an exemplary method 200 of producing a consistent and potentially unlimited supply 202 of a reference nucleic acid. In one implementation, the exemplary method 200 includes synthesizing a tagged target reference nucleic acid 204 to be amplified. The (untagged) reference nucleic acid 100 to be amplified, of course, has a base sequence capable of being used at least in part as a clinical reference. The exemplary method 200 includes synthesizing a first exemplary tag 206 on one end of the reference nucleic acid 100 to be amplified, for example, oriented for “forward” sequence extension, and a second exemplary tag 208 on the opposing end of the reference nucleic acid 100 to be amplified, for example, oriented for “reverse” sequence extension of a complementary strand. In one implementation, the first exemplary tag 206 has a base sequence complementary to a base sequence of a first primer of a primer set and the second exemplary tag 208 has the same base sequence as a second primer of the primer set. The first tag 206 anneals to one of the primers (or vice versa

depending on viewpoint) and the resulting primed strand becomes a template for producing its complement strand through extension of the first primer. The second tag 208 anneals to no primer, but provides a template for making a priming site on the complementary strand. Thus the complementary strand includes a complement of the second tag 208, which anneals to the second primer. The primed complement strand becomes a template for producing a copy of the original nucleic acid strand by extension of the second primer. Both the original tagged reference nucleic acid 204 and its complementary strand can be amplified at the same time during PCR cycles.

A potentially unlimited supply of a consistently reproducible reference nucleic acid 202 can thus be obtained for use in a myriad of clinical tests (e.g., 212, 214, 216, ..., 218) where the easy availability of supply and the reproducibility of the reference nucleic acids are desirable as features for a standard to be used for comparing results between different types of clinical tests.

Fig. 3 shows another exemplary method 300 according to another aspect of the subject matter. A starting mixture 302 includes multiple reference nucleic acid constituents each tagged with copies of the same exemplary first and second nucleic acid tags (e.g., 206, 208 in Fig. 2). The starting mixture 302 undergoes a PCR process, such as an exemplary mixed-template PCR 304, to yield a supply of consistently reproducible reference nucleic acids in a single mixture 306.

When the starting mixture 302 includes constituent reference nucleic acids that are oligonucleotides or polynucleotides each synthesized to represent a different standard reference, then the exemplary method 300 produces a single mixture 306 that can be used as a single universal

reference for many different types of tests, that is, as a multiplexed assay or universal reference, at least with respect to tests compatible with the particular type of constituent reference nucleic acids in the single mixture 306 being produced (in other words, the universality of a single mixture 306 may be limited by the reagent characteristics specifically required by specific conventional machines or other conventional apparatuses). An exemplary system for utilizing reference nucleic acids in a single mixture 306 is discussed below with respect to Figure 11.

An exemplary supply of reference nucleic acids in a single mixture 306 may replace multiple separate conventional nucleic acid references. For example, cystic fibrosis is a disease that may be caused by the mutation of a single base pair in any one of several hundred different sites on the cystic fibrosis transmembrane conductance regulator (CFTR) gene. Testing for 30-33 of the commonest of these mutations—using 30-33 different nucleic acid reference “controls”—can provide a successful diagnosis for over 90% of cystic fibrosis cases in Caucasian populations. Instead of requiring the conventional 30-33 different nucleic acid references in 30-33 different containers (or cumbersome mixtures thereof), an exemplary method 300 using the subject matter produces a supply of reference nucleic acids in a single mixture 306 for all 30-33 tests: the single mixture 306 is amplified as a mixture. This exemplary method can be scaled to produce a supply of much more than just 30-33 different nucleic acid references in a single mixture 306.

Exemplary Synthetic Reference Nucleic Acids

Fig. 4 shows, in greater detail than in Fig. 2, an original tagged reference nucleic acid 204 (SEQ ID NO: 1 plus SEQ ID NO: 2 plus SEQ ID

NO: 3) to be used as a starting template in a PCR process. The (untagged) reference nucleic acid 100 SEQ ID NO: 1 to be amplified is typically a segment of a larger nucleic acid polynucleotide, such as a genomic nucleic acid 400 that has a base sequence at least a part of which is used as a clinical reference or as a building block for a clinical reference. The base sequence and length of the illustrated reference nucleic acid 100 SEQ ID NO: 1 to be amplified is merely an example for description purposes. Actual examples of oligonucleotides and/or polynucleotides to be amplified may be different in sequence and length than those illustrated. Likewise, the first primer 404 SEQ ID NO: 4 and second primer 406 SEQ ID NO: 5 are depicted as a universal M13 forward primer (that has seventeen bases) and a universal M13 reverse primer (that has sixteen bases), respectively, only for the sake of example. Other primers may be used or designed, such as any one of an M13 phage vector sequence or primer, Lambda, SP6, T3, or T7 primers, or any arbitrarily designed base sequence present in nature or completely invented by the architect, depending on, or independent of the reference nucleic acid 100 SEQ ID NO: 1 to be amplified.

In one implementation, the base sequence of each tag (206, 208 – SEQ ID NO: 2 and SEQ ID NO: 3) is typically designed to be unique so that the bases in the tags do not pair with a span of the bases in the reference nucleic acid 100 SEQ ID NO: 1 to be amplified (or complementary strands thereof). Using tags (206, 208 – SEQ ID NO: 2 and SEQ ID NO: 3) with unique base sequences prevents formation of unwanted artifacts, such as secondary nucleic acid structures, duplexes, concatemers, etc. Each tag is bound to its oligonucleotide in an orientation that allows an associated primer to extend copies of the original strand and complementary strands.

In another implementation, more than one pair of tags may be synthesized to allow more flexibility for manipulating a specific nucleic acid sequence or set of sequences in a complex mixture of sequences. The multiple pairs of tags may be attached to selected reference nucleic acid species in the mixture to provide differential propagation and control of the different selections. Hence, it is possible to orchestrate propagation of very complex reference nucleic acid mixtures having precise proportions of the constituents, e.g., by varying multiple primer and/or target nucleic acid concentrations.

10 In the illustrated implementation, the first tag 206 SEQ ID NO: 2 has a base sequence complementary to a first primer 404 SEQ ID NO: 4 of a primer set 402. The complementariness between the first tag 206 SEQ ID NO: 2 and the first primer 404 SEQ ID NO: 4 allows the first primer 404 SEQ ID NO: 4 to anneal to the original tagged reference nucleic acid 204
15 (SEQ ID NO: 1 plus SEQ ID NO: 2 plus SEQ ID NO: 3) and extend (that is, be extended by a polymerase) in the 3' direction—new bases are added by the polymerase on to the 3' end of the primer.

The second tag 208 SEQ ID NO: 3 has a base sequence that matches (i.e., is the same as) a base sequence of a second primer 406 SEQ ID NO: 5
20 of the primer set 402. The second primer 406 SEQ ID NO: 5 does not anneal to the original tagged reference nucleic acid 204 since its sequence is the same as the second tag 208 SEQ ID NO: 3 segment of the original tagged reference nucleic acid 204, but anneals to a complementary priming site (506 in Fig. 5 -- SEQ ID NO: 8) created on a complementary nucleic acid strand
25 by extension of the first primer 404 SEQ ID NO: 4, for example, in a previous PCR cycle. Hence, on an original tagged reference nucleic acid

204, the first tag 206 SEQ ID NO: 2 acts as a priming site while the second tag 208 SEQ ID NO: 3 creates a priming site on the strand complementary to the original tagged reference nucleic acid 204.

Fig. 5 shows an exemplary reference nucleic acid production method

5 500 according to the subject matter in which a first primer 404 SEQ ID NO: 4 priming a synthetic tagged reference nucleic acid 204 (SEQ ID NO: 1 plus SEQ ID NO: 2 plus SEQ ID NO: 3) is extended into a complementary strand 502 (SEQ ID NO: 4 plus SEQ ID NO: 6) of the original synthetic tagged reference nucleic acid 204, thus forming a duplex (204 bound to 502). The

10 complementary strand 502 includes a priming site 506 SEQ ID NO: 8 that is complementary to the second tag 208 SEQ ID NO: 3. After the duplex undergoes denaturation, a second primer 406 SEQ ID NO: 5 priming the complementary strand 502 is extended into a copy of the original tagged reference nucleic acid 204. The amplification of the original tagged

15 reference nucleic acid 204 and its complement [[506]] 502 continues through multiple PCR cycles using, in one implementation, exemplary PCR components and reaction parameters, such as those indicated in Appendix A: “Synthetic Control Design Parameters: Exemplary PCR Components and Reaction Parameters,” incorporated herein by reference. The duplex is

20 denatured as per known PCR methods to separate the strands (204, [[506]] 502) of the duplex. In the annealing and extension parts of subsequent PCR cycles, the second primer 406 SEQ ID NO: 5 makes further copies of the original tagged reference nucleic acid 204 while the first primer 404 SEQ ID NO: 4 makes further copies of complementary strands of the original tagged

25 reference nucleic acid 204. The PCR cycles of denaturation, annealing, extension, and denaturation are repeated, e.g., approximately thirty times, to

produce millions of copies of the original tagged reference nucleic acid 204 having the reference base sequence for clinical use and millions of copies of its complementary strand 502.

Fig. 6 shows an exemplary synthetic reference nucleic acid synthesized as a duplex 600, that is, as two complementary nucleic acid oligonucleotides 602 (SEQ ID NO: 7 plus SEQ ID NO: 8), and 604 (SEQ ID NO: 9 plus SEQ ID NO: 2), or, polynucleotides each having one exemplary tag on one end). This provides an alternative exemplary method for synthesizing reference nucleic acids that include or do not include mutations, using one or more of the exemplary tags described with respect to Figs. 4 and 5. Hence, a single-stranded reference nucleic acid can be synthesized to include a tag on each end (Figs. 4 and 5), which thereafter relies on PCR reagents and reactions to produce complements, or two complementary strands, each having one tag, can be synthesized from the outset (as in Fig. 6).

Exemplary Mixtures of Synthetic Reference Nucleic Acids

Fig. 7 shows an exemplary mixture 701 of tagged reference nucleic acids (702, 704, 706, 708, 710, 712, ..., 714) that may be used as a starting material for a PCR amplification process to obtain a potentially unlimited, reproducible, and consistent supply of reference nucleic acid mixture 718. Copies of a first nucleic acid tag 206 are bound to a first end of each of the multiple reference nucleic acids, wherein each copy of the first nucleic acid tag 206 has a base sequence complementary to a base sequence of a first primer 404 of a primer set 402. Copies of a second nucleic acid tag 208 are bound to a second end of each of the multiple reference nucleic acids,

wherein each copy of the second nucleic acid tag 208 has a base sequence matching a base sequence of a second primer 406 of the primer set 402.

During an exemplary mixed-template PCR 716, a single primer set 402 is added to the exemplary mixture 701 of tagged reference nucleic acids.

5 The exemplary mixed-template PCR 716 produces a supply of reference nucleic acids mixture 718 that can be reproduced consistently in potentially unlimited quantity.

In one implementation, the tagged reference nucleic acids (702, 704, 706, 708, 710, 712, ..., 714) are each synthesized to represent at least one
10 mutation in a base sequence of a gene, for example, the CFTR gene. The different tagged reference nucleic acids (702, 704, 706, 708, 710, 712, ..., 714) are designed and synthesized, when possible, to have similar lengths if such similarity facilitates uniform amplification.

Table 1 below is a list of known common mutations of the CFTR
15 gene causing cystic fibrosis (Cystic Fibrosis Genetic Analysis Consortium (1994)). Although the subject matter can be used with respect to many different types of genetic testing, cystic fibrosis mutation testing is used as an illustrative disease. One example of a consistent supply of a reference nucleic acid mixture 718 that is designed, synthesized, and propagated
20 according to the subject matter includes a tagged reference nucleic acid for each CFTR gene mutation listed below in Table 1, that is, for each mutation listed in Table 1 a tagged reference nucleic acid is synthesized with a base sequence that includes [[and]] at least one, and possibly more than one of the known mutation:

Table 1

Name of Mutation
[[Delta]]F508
G542X
G551D
N1303K
W1282X
R553X
621+1G->T
1717-1G->A
R117H
R1162X
R347P
3849+10kbC->T
[[Delta]]I507
394delTT
G85E
R560T
A455E
1078delT
2789+5G->A
3659delC
R334W
1898+1G->T
711+1G->T
2183AA->G
3905insT
S549N
2184delA
Q359K/T360K
M1101K
Y122X
1898+5G->T
3120+1G->A
I148T

Of course, other exemplary supplies of reference nucleic acid mixtures 718 designed, synthesized, and/or propagated according to the subject matter may have different constituent tagged reference nucleic acids comprising more CFTR gene mutations or fewer CFTR gene mutations than those shown in Table 1. Still other exemplary supplies of reference nucleic acids mixtures 718 designed, synthesized, and/or propagated according to the subject matter may have different tagged reference nucleic acids used in tests for other genetic conditions and circumstances unrelated to cystic fibrosis.

Fig. 8 shows an exemplary complex mixture 800 (also known as “master mixture”) of reference nucleic acid sequences 800, i.e., a “mixture of mixtures” in which more than one pair of tags may be synthesized to allow more flexibility for manipulating a specific nucleic acid sequence or set of sequences in the complex mixture 800. As mentioned, the multiple pairs of tags may be attached to selected reference nucleic acid species in the mixture to provide differential propagation and control of the different selections. Hence, it is possible to orchestrate propagation of very complex reference nucleic acid mixtures having precise proportions of the constituents, e.g., by varying multiple primer concentrations and/or by strategically attaching various tag sets to predetermined reference nucleic acid sequences.

In one implementation, each member of a first mixture of diverse reference nucleic acids (702, 704, 706, 708, 710) may be flanked by copies of the same set of tags 796, 798 and may be controlled accordingly via a primer set tuned to the tags 796, 798 (e.g., M13 tuned tags). Each member of a second mixture of diverse reference nucleic acids (802, 804, 806, 808, 810) may be flanked by copies of a different set of tags 896, 898 (e.g., T3/T7

tuned tags) and may be controlled accordingly via a second primer set tuned to the tags 896, 898. The process of selecting different tags to control different mixtures within one “master” mixture 800 may be extended to multiple mixtures and/or multiple individual species within the master
 5 mixture 800.

In the same or another implementation, by adding specific amounts of the primers that are specific for tagged sequences that require special treatment the level of amplification and other enzymatic manipulation can be finely adjusted. Hence, if a reference nucleic acid sequence “X” is
 10 particularly difficult to visualize in a detection assay, it may be desirable to have ten times more of sequence X than any other. Sequence X can be tagged with unique tags (e.g., TX1 and TX2) whereas other nonremarkable reference nucleic acid sequences, A, B, C, . . . Z are all tagged with the same tags (e.g., TA1 and TA2), but different tags than used for sequence X. This
 15 provides more control over manipulating sequence X than for controlled the other sequences in the complex mixture 800.

Fig. 9 shows an exemplary reference nucleic acid mixture 900, wherein each constituent reference nucleic acid (902, 904, 906, 908) embodies additional design features for producing a mixture product of
 20 consistent and reproducible quality. Each single mutation represented in each different reference nucleic acid may be surrounded by base intervals 910, 912 of similar length, that is, each mutation site can be surrounded on both sides by approximately the same number of bases between the site of a mutation and each nucleic acid tag 206 SEQ ID NO: 10 and 208 SEQ ID
 25 NO: 11. This “padding” of roughly uniform length between a mutation site and tags on a given reference nucleic acid distances the mutation site from

possible interference with PCR processes, other subsequent enzymatic manipulation, and other processes involving the tags [[[]] 206 SEQ ID NO: 10 and 208 [[]] SEQ ID NO: 11]. The base intervals 910, 912 around a mutation site can be synthesized using the same base sequence arrangements as are present at corresponding sites of a normal model gene on which the mutation occurs, if such a gene exists, or can be synthesized as arbitrary sequences. In one implementation, the similarly-sized base intervals 910, 912 on either side of one or more mutations may add up to a span of between approximately forty and approximately sixty bases, resulting in total polynucleotide lengths that provides a good value given synthesis costs. In other implementations, the base intervals 910, 912 on either side of one or more mutation sites may add up to spans of approximately eighty, one hundred, or even more than one hundred bases.

Likewise, as shown in **Fig. 10**, if two or more mutations are present in an exemplary ~~tagged~~ reference nucleic acid 1000 with tags on each end (SEQ ID NO: 10 and SEQ ID NO: 11), substantially uniform base intervals 910, 912, 1002 may be designed into the reference nucleic acid 1000 to separate the two mutations. Thus, design geometry for optimizing PCR amplification and subsequent manipulation of the PCR product in light of multiple different references being present in a mixture can be implemented in each of multiple reference nucleic acids in a mixture, as exemplified by the illustrated reference nucleic acid 1000. Design considerations for references to be used in a mixture of references can include those that prevent steric impediments to PCR amplification and subsequent manipulation of the PCR product, such as avoiding a variety of significantly different sized reference nucleic acids as starting templates and avoiding tag

designs that might lead to secondary structure artifacts in a mixture of references.

Since each constituent reference nucleic acid (e.g., 1000) to be used as an original tagged reference nucleic acid 204 in a mixture can be synthesized de novo by the subject matter, problem configurations can be designed out of references. The resulting consistent reference nucleic acid supply (e.g., 718) can therefore achieve superior quality, reproducibility and consistency over conventional methods.

10 Exemplary Reference Nucleic Acid Extension Methods

Fig. 11 shows an exemplary ligation extension method 1100 that can be used alone or in addition to other exemplary techniques described herein. The exemplary method 1100 can join two oligonucleotides, allowing a reference nucleic acid architect to design and synthesize smaller, simpler oligonucleotides to be joined into longer, more complex oligo- or polynucleotides later on in the design, synthesis, amplification cycle—before (or in some cases after) amplification. The exemplary method 1100 provides another tool for designing, synthesizing, and/or propagating reference nucleic acids, in this case a tool for facilitating rapid synthesis of longer nucleic acid base sequences for use as references in molecular diagnostic and genetic testing of human or nonhuman nucleic acid base sequences. In other words, the exemplary method 1100 allows a larger reference nucleic acid to be built from smaller, more manageable, and more manipulable pieces.

In one implementation, the exemplary method 1100 uses exemplary ligation extension components and reaction parameters, such as those indicated in Appendix B: “Synthetic Control Design Parameters: Exemplary

Ligation Extension Components and Reaction Parameters,” incorporated herein by reference.

An exemplary method 1100 can be used for joining two of any amenable types of oligonucleotides, but is especially useful for joining synthetic oligonucleotides e.g., 1102[[:]] SEQ ID NO: 12 and 1104 SEQ ID NO: 13 that represent adjacent segments of a genomic nucleic acid 400, wherein the adjacent segments modeled by the synthetic oligonucleotides 1102 SEQ ID NO: 12 and 1104 SEQ ID NO: 13 are both relevant as clinical references. Alternatively, it may be desirable to join two reference nucleic acids for physical, chemical, steric, and/or practical reasons related to supply and packaging logistics of a particular reference nucleic acid mixture 718.

Each oligonucleotide to be ligated (e.g., 1102 SEQ ID NO: 12 and 1104 SEQ ID NO: 13) is synthesized with an exemplary tag (e.g., 206 SEQ ID NO: 2 and 208 SEQ ID NO: 5) on the end opposite of its candidate ligation site. An additional oligonucleotide comprising a complement bridge sequence 1106 SEQ ID NO: 14 is introduced accompanied by ligase (a ligation enzyme) and dNTPs (the individual base unit “monomers” to become the building blocks for strand extension). One end of the complement bridge sequence 1106 SEQ ID NO: 14 has a base sequence complementary to the non-primer-tag end of one of the oligonucleotides 1102 SEQ ID NO: 12 to be ligated and the other end of the complement bridge sequence 1106 SEQ ID NO: 14 has a base sequence complementary to the non-primer-tag end of the other oligonucleotide 1104 SEQ ID NO: 13 to be ligated. The tagged reference oligonucleotides—1102 bound to 208[[:]] (SEQ ID NO: 12 plus SEQ ID NO: 5) and 1104 bound to 206 (SEQ ID NO: 13 plus SEQ ID NO: 2) align by complementation onto the

complement bridge sequence 1106 SEQ ID NO: 14 that acts as template and “bridge.” The ligase joins the two or more oligonucleotides (e.g., 1102 SEQ ID NO: 12 and 1104 SEQ ID NO: 13), creating a single oligonucleotide 1108 (SEQ ID NO: 15 plus SEQ ID NO: 2 plus SEQ ID NO: 5) from the two
5 original oligonucleotides 1102 SEQ ID NO: 12 and 1104 SEQ ID NO: 13.

Once the single oligonucleotide 1108 is created from the two original oligonucleotides 1102, 1104, the single oligonucleotide 1108 may be amplified singly or in an exemplary mixture 601 by an exemplary method such as those depicted with regard to Figs. 5 and 6.

10 **Fig. 12** shows an exemplary overlap extension method 1200 that can be used alone or in addition to other exemplary techniques described herein. The exemplary method 1200 can join two or more oligonucleotides 1102 SEQ ID NO: 12 and 1104 SEQ ID NO: 13 on either side of a region of complementation 1202, allowing (in a manner similar to the exemplary
15 method 1100 above) the reference nucleic acid architect to design and synthesize smaller, simpler oligonucleotides to be joined into longer, more complex oligonucleotides or polynucleotides later on in the design, synthesis, amplification cycle—i.e., before (or in some cases after) amplification. The exemplary method 1200 provides another member in a
20 toolbox of exemplary techniques for designing, synthesizing, and/or propagating reference nucleic acids for use as references in molecular diagnostic and genetic testing of human or nonhuman nucleic acid base sequences. The exemplary method 1200 allows a larger reference nucleic acid to be built from smaller, more manageable, and more manipulable
25 pieces.

In one implementation, the exemplary method 1200 uses exemplary overlap extension components and reaction parameters, such as those indicated in Appendix C: “Synthetic Control Design Parameters: Exemplary Overlap Extension Components and Reaction Parameters,” incorporated
5 herein by reference.

In the exemplary overlap extension method 1200, a region of complementation 1202 is designed into the base sequence of one or both of the oligonucleotides 1102, 1104 to be overlapped, except in the unlikely event that both oligonucleotides 1102, 1104 already have complementary
10 ends opposite to their respective tag ends (i.e., the ends bearing tags associated with a primer set [[402]] for amplification). In the illustrated case, an end sequence SEQ ID NO: 16 is synthesized onto oligonucleotide 1102 SEQ ID NO: 12 to provide complementation with part of oligonucleotide 1104 SEQ ID NO: 13.

15 So that the duplex 1204 resulting from this exemplary method 1200 may be amplified singly or in an exemplary mixture 601 by an exemplary method such as those depicted with regard to Figs. 5 and 6, tags (e.g., 206 SEQ ID NO: 2 and 506 SEQ ID NO: 8) for use with a primer set (e.g., 402) are designed and synthesized onto the not-to-be-overlapped ends of each of
20 the reference oligonucleotides 1102, 1104, that is, on the ends opposite to the respective regions of complementation 1202.

In one implementation, the tags (e.g., 206 SEQ ID NO: 2 and 506 SEQ ID NO: 8) for both reference oligonucleotides 1102, 1104 are complementary to their respective primers (e.g., 404 SEQ ID NO: 4 and 406
25 SEQ ID NO: 5). It should be noted that multiple reference oligonucleotides can intervene and be incorporated by exemplary ligation techniques between

the end reference nucleotides 1102 and 1104 bearing the exemplary tags complementary to the primers, as will be discussed with respect to Fig. 13. When the oligonucleotides 1102, 1104 bound to their corresponding regions of complementation 1202 and to their tags 205, 506 are placed together with

5 Taq polymerase and dNTPs, but no primers for PCR, the corresponding one or more regions of complementation 1202 anneal to each other and each act as a primer for extending the two or more overlapped stands (1102 bound to both 506 and 1202, overlapping 1104 bound to 206) in opposite directions (1208, 1210) into a duplex 1204 longer than either of the overlapped strands,

10 consisting of example longer polynucleotide strands 1214 SEQ ID NO: 17 and 1216 SEQ ID NO: 18.

Each end strand (i.e., bearing a tag) of the duplex 1204 receives a respective end base sequence 1212 SEQ ID NO: 4 and 208 SEQ ID NO: 5 that matches the base sequence of respective members of the primer set 402

15 (sequences 404 SEQ ID NO: 4 and 406 SEQ ID NO: 5). Each end strand already includes a tag (206 SEQ ID NO: 2 and 506 SEQ ID NO: 8) complementary to the remaining member of the primer set 402 (sequences 404 SEQ ID NO: 4 and 406 SEQ ID NO: 5). Hence, each end strand of the resulting duplex 1204 includes the correct tags for being amplified singly or

20 in an exemplary mixture 601 by an exemplary method such as those described with regard to Figs. 5 and 6 and can enter into those exemplary methods 500, 600 using a denaturation step to separate the strands.

Fig. 13 shows exemplary ligation methods 1300 in which multiple complement bridge sequences 1106 (not necessarily identical) and/or

25 multiple regions of complementation 1202 (not necessarily identical) are used to ligate multiple synthesized reference nucleic acid segments 1302.

Thus, a single reference nucleic acid may be synthesized by ligating the multiple segments 1302, by exemplary methods described with respect to Figs. 11 and/or 12, wherein end sequences 1304, 1306; 1308, 1310 bear exemplary tags.

5 **Fig. 14** shows an exemplary method 1400 of designing, synthesizing, and/or propagating one or more reference nucleic acids for carrier testing: to simulate at least part of the genetic profile of a heterozygous carrier. The resulting reference nucleic acid or mixture of reference nucleic acids may be synthetic or semi-synthetic.

10 In one implementation, a mixture of different pairs of homologous nucleic acid segments 1402 (one pair illustrated) are designed, synthesized, and propagated using one or more of the exemplary methods described herein. That is to say, for at least some pieces of reference nucleic acid designed and synthesized with one or more mutations 1404 by an exemplary
15 method, homologous pieces of nucleic acid 1406 without the one or more mutations are also designed and synthesized. This implementation of the exemplary method 1400 can produce a mixture of reference nucleic acids that is heterozygous at particular loci wherein a mutation occurs in only one member of the homologous pair, e.g., in only one of the alleles of a
20 homologous gene pair. This simulates the heterozygous makeup of a heterozygous carrier.

 In another implementation of an exemplary method 1400, homozygotes that include one or more mutations 1408 are designed and synthesized by one or more of the exemplary methods described herein.
25 Normal human genomic DNA (or other normal human or animal or plant of other organism nucleic acid) is added to the mixture, thereby forming

heterozygous loci wherever there is a piece of homologous reference nucleic acid bearing a mutation. Such an exemplary mixture of synthetic elements and human biological elements simulates homologous gene pairs that are heterozygous at every locus of interest—in other words, the exemplary
5 mixture provides carrier testing across many different loci.

To reiterate, in one implementation, an exemplary mixture of synthesized multiple reference nucleic acids for carrier testing is designed and synthesized using exemplary tags tuned to primer sets as described herein, and the exemplary mixture also contains partial or total genomic
10 DNA from a reference human source that is known to differ at least one locus from one or more of the synthesized multiple reference nucleic acids.

Exemplary Systems

Fig. 15 shows an exemplary universal genetic mutation testing system
15 (UGMTS) 1500. Such an exemplary UGMTS 1500 includes hardware, software, lab ware, chemical processors, chemical reagents, control logic, etc., compatible with using an exemplary reference nucleic acid mixture 718 produced as described herein as a universal reference or control, wherein the universe of tests for which the reference nucleic acid mixture 718 is
20 “universal” typically includes a battery of genetic mutation tests associated with a certain gene or a particular disease.

In one implementation, the reference nucleic acid mixture 718 includes multiple constituent reference nucleic acids that are normal manifestations of patient gene sites being tested, and therefore the reference
25 nucleic acid mixture 718 is compared directly with the patient’s nucleic acids. In another implementation, the reference nucleic acid mixture 718

includes multiple constituent reference nucleic acids that are manifestations of mutated gene sites being tested, and therefore the reference nucleic acid mixture 718 is used to calibrate and/or verify a set of normal gene controls, which are being compared directly with the patient's nucleic acid, or to
5 calibrate and/or verify the operation of the apparatus (1500) itself.

In an example configuration of the former implementation, an exemplary reference nucleic acid mixture 718 has been designed, synthesized, and amplified so that it contains constituent reference nucleic acids that may be compared directly with amplified samples of a patient's
10 nucleic acid in order to perform a battery of tests for different genetic mutations that could be causing the same disease. Accordingly, an exemplary UGMTS 1500 includes a patient sample input 1502 coupled with a target gene amplifier 1504. The target gene amplifier 1504 may be a PCR apparatus or PCR process to produce a large enough quantity of a patient's
15 target gene (gene of interest) to perform a battery of mutation tests.

The reference nucleic acid mixture 718 and the patient's nucleic acid from the target gene amplifier 1504 are both fed to an electrophoresis module 1506 that includes a gel box 1508 for simultaneous gel electrophoresis of the patient's nucleic acid and the reference nucleic acid
20 mixture 718. The gel box 1508 is supported by a running buffer supply 1510 and a capillary electrophoresis module 1512. A spectrophotometry module 1514 coupled with the electrophoresis module 1506 includes a laser scanner 1516 to read the results of the gel electrophoresis. A comparator 1518 evaluates readings obtained by the laser scanner 1516 for presentation on a
25 display 1520. The described UGMTS 1500 is supported by appropriate power supplies 1522 and control logic 1524.

The exemplary UGMTS 1500 described above is only one example system that compatibly uses an exemplary reference nucleic acid mixture 718 produced according to the subject matter. Other exemplary systems that include different parts and/or other configurations can also use an exemplary
5 reference nucleic acid mixture 718.

CONCLUSION

The foregoing describes exemplary reference nucleic acids, mixtures of reference nucleic acids, methods of designing, synthesizing, and
10 propagating each, and an exemplary system for using a reference nucleic acid mixture. The description is not meant to be limiting but illustrative of the subject matter.

CLAIMS

1. A method, comprising:
synthesizing one or more nucleic acid sequences relevant for clinical
5 reference;
tagging at least one end of each sequence for amplification by a
primer; and
amplifying the one or more nucleic acid sequences using the primer.
- 10 2. The method as recited in claim 1, wherein the tagging includes
attaching an additional sequence of nucleotides, wherein the additional
sequence is complementary to a nucleotide sequence of the primer.
3. The method as recited in claim 1, wherein the tagging includes
15 attaching an additional sequence of nucleotides, wherein the additional
sequence is identical to a nucleotide sequence of the primer.
4. The method as recited in claim 1, wherein:
the tagging includes attaching a first sequence of nucleotides to a first
20 end of each of the one or more synthesized nucleic acid sequences, wherein
the first sequence is complementary to a nucleotide sequence of a first
primer of a primer set, and
the tagging includes attaching a second sequence of nucleotides to a
second end of each of the one or more synthesized nucleic acid sequences,
25 wherein the second sequence is identical to a nucleotide sequence of a
second primer of a primer set.

5. The method as recited in claim 1, wherein the synthesizing comprises synthesizing two complementary nucleic acid strands, wherein:

5 a first strand includes a first nucleic acid sequence relevant for clinical reference and a nucleic acid tag complementary to a first primer of a primer set and

a second strand includes a nucleic acid sequence complementary to the first strand and a nucleic acid tag complementary to a second primer of a primer set.

10

6. The method as recited in claim 5, wherein the first and second strands are synthesized as a duplex of complementary strands.

7. The method as recited in claim 1, wherein at least one of the
15 one or more synthesized nucleic acid sequences includes at least one mutation of a nucleotide in a normal human nucleic acid.

8. The method as recited in claim 1, further comprising synthesizing multiple mixtures of at least one reference nucleic acid apiece,
20 wherein:

each of the multiple mixtures has an associated primer set, and wherein:

each member of one of the multiple mixtures includes a first tag attached to a first end of the member, wherein:

the first tag comprises a sequence of nucleotides complementary to a nucleotide sequence of a first primer of the associated primer set, and

each member includes a second tag attached to a second end of
5 the member, wherein:

the second tag comprises a sequence of nucleotides identical to a nucleotide sequence of a second primer of the associated primer set.

10 9. The method as recited in claim 8, further comprising combining the multiple mixtures and controlling each of the multiple mixtures to achieve separate amounts of amplification for each of the multiple mixtures.

15 10. The method as recited in claim 9, wherein controlling each of the multiple mixtures includes controlling an amount of each associated primer set for each of the multiple mixtures.

20 11. The method as recited in claim 9, wherein controlling each of the multiple mixtures includes controlling a physical characteristic of a combined mixture of the multiple mixtures to favor an amplification capability of one primer set over an amplification capability another primer set.

25 12. The method as recited in claim 9, wherein controlling each of the multiple mixtures includes controlling a physical characteristic of a PCR

reaction to favor an amplification capability of one primer set over an amplification capability another primer set.

13. The method as recited in claim 9, wherein one of the multiple
5 mixtures includes only one reference nucleic acid.

14. The method as recited in claim 1, wherein the synthesizing
further comprises synthesizing a nucleic acid that includes a normal human
base sequence and synthesizing a corresponding copy of the nucleic acid that
10 includes at least one mutation of the normal human base sequence.

15. The method as recited in claim 14, further comprising using an
amplified mixture of both the nucleic acid that includes a normal human base
sequence and the corresponding copy that includes at least one mutation to
15 more accurately simulate a heterozygous state.

16. The method as recited in claim 1, further comprising adding
normal human nucleic acid to the one or more synthesized nucleic acid
sequences relevant for clinical reference in order to achieve a mixture of the
20 nucleic acids representing at least a segment of homologous heterozygous
alleles.

17. A method, comprising:
synthesizing a mixture of nucleic acids, wherein each nucleic acid is
25 capable of being used as a reference in mutation testing and wherein a first
end of each nucleic acid in the mixture includes a base sequence

complementary to a base sequence of a first primer and a second end of each nucleic acid in the mixture includes a base sequence matching a base sequence of a second primer; and

amplifying the mixture of nucleic acids during at least one polymerase chain reaction using the first primer and the second primer.

18. The method as recited in claim 17, wherein at least one of the nucleic acids is capable of being used as a reference by including at least one mutation of a normal base sequence.

10

19. The method as recited in claim 17, further comprising amplifying the mixture of nucleic acids through multiple polymerase chain reaction cycles to produce an ongoing supply of the mixture.

15 20. The method as recited in claim 17, further comprising using the mixture of nucleic acids as at least one reference in a mutation test.

21. The method as recited in claim 17, further comprising joining multiple nucleic acid segments using a ligation extension to make at least one of the reference nucleic acids in the mixture.

20

22. The method as recited in claim 21, wherein for at least one of the reference nucleic acids, the synthesizing includes:

synthesizing a first nucleic acid that includes a first end comprising a base sequence complementary to the base sequence of the first primer and a

25

second end complementary to a base sequence on a first end of a bridge nucleic acid;

synthesizing a second nucleic acid that includes a first end comprising a base sequence that matches the base sequence of the second primer and a
5 second end complementary to a second end of the bridge nucleic acid; and

making the reference nucleic acid by joining multiple nucleic acid segments in the ligation extension, including joining the first nucleic acid on one end of the joined segments using the bridge nucleic acid and joining the second nucleic acid on the opposite end of the joined segments using the
10 bridge nucleic acid.

23. The method as recited in claim 17, further comprising joining multiple nucleic acids using an overlap extension to make at least one of the reference nucleic acids in the mixture.

15

24. The method as recited in claim 23, wherein for at least one of the reference nucleic acids, the synthesizing includes:

synthesizing a first nucleic acid that includes a first end comprising a base sequence complementary to the base sequence of the first primer;

20 synthesizing a second nucleic acid that includes a first end comprising a base sequence complementary to the base sequence of the second primer and a second end complementary to an end of a nucleic acid segment to be adjacent to the second nucleic acid in the final reference nucleic acid; and

making the final reference nucleic acid by joining multiple nucleic acid segments in overlap extensions, including joining the first nucleic acid
25 on one end of the joined segments using the overlap extension and joining

the second nucleic acid on the opposite end of the joined segments using the overlap extension in the absence of the first and second primers.

25. A method of propagating a reference nucleic acid, comprising:
5 synthesizing a single-stranded reference nucleic acid;
tagging the reference nucleic acid to make a tagged reference nucleic acid, comprising:

synthesizing a first nucleic acid tag bound to a first end of the reference nucleic acid, wherein the first nucleic acid tag has a base sequence
10 complementary to a base sequence of a first primer of a primer set and

synthesizing a second nucleic acid tag bound to a second end of the reference nucleic acid, wherein the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set; and
subjecting the tagged reference nucleic acid to a polymerase chain
15 reaction in the presence of the first primer and the second primer.

26. The method as recited in claim 25, wherein the first primer produces copies of a single-stranded nucleic acid complementary to the tagged reference nucleic acid and the second primer produces copies of the
20 tagged reference nucleic acid.

27. The method as recited in claim 25, wherein the first nucleic acid tag is bound to the 3' end of the reference nucleic acid and the second nucleic acid tag is bound to the 5' end of the reference nucleic acid.

28. The method as recited in claim 25, wherein the synthesizing and the tagging include a ligation extension of two or more nucleic acids.

29. The method as recited in claim 25, wherein the synthesizing
5 and the tagging include an overlap extension of two or more nucleic acids.

30. A method of propagating a mixture of reference nucleic acids, comprising:

synthesizing multiple single-stranded reference nucleic acids;

10 tagging each of the multiple reference nucleic acids to make tagged reference nucleic acids, comprising:

synthesizing a copy of a first nucleic acid tag bound to a first end of each of the multiple reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence
15 of a first primer of a primer set and

synthesizing a copy of a second nucleic acid tag bound to a second end of each of the multiple reference nucleic acids, wherein each copy of the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set;

20 mixing the multiple tagged reference nucleic acids with each other; and

subjecting the mixture of tagged reference nucleic acids to a polymerase chain reaction in the presence of the first primer and the second primer.

31. The method as recited in claim 30, wherein the first primer produces single-stranded nucleic acids complementary to each of the multiple tagged reference nucleic acids and the second primer produces copies of each of the multiple tagged reference nucleic acids.

5

32. The method as recited in claim 30, wherein each copy of the first nucleic acid tag is bound to the 3' end of each of the multiple reference nucleic acids and each copy of the second nucleic acid tag is bound to the 5' end of each of the multiple reference nucleic acids.

10

33. The method as recited in claim 30, wherein each of multiple reference nucleic acids in the mixture of reference nucleic acids include at least a base sequence that represents a different mutation of at least part of a gene.

15

34. The method as recited in claim 33, wherein the mixture is used in mutation testing.

35. The method as recited in claim 34, wherein the mutation testing comprises testing a base sequence from a cystic fibrosis transmembrane conductance regulator gene.

20

36. A method of synthesizing a tagged reference nucleic acid for propagation, comprising:

25 synthesizing a single-stranded reference nucleic acid;

synthesizing a first nucleic acid tag bound to a first end of the reference nucleic acid, wherein the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer set.

5 37. The method of synthesizing as recited in claim 36, further comprising synthesizing two of the single-stranded reference nucleic acids, wherein the two are complementary.

10 38. The method as recited in claim 36, wherein a single-stranded reference nucleic acid includes a mutation of a normal human nucleic acid sequence.

15 39. The method as recited in claim 36, further including synthesizing a second nucleic acid tag bound to a second end of the reference nucleic acid, wherein the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

20 40. The method as recited in claim 36, wherein the reference nucleic acid includes a base sequence representing a mutation of a cystic fibrosis transmembrane conductance regulator gene sequence for cystic fibrosis mutation testing.

25 41. A method of synthesizing a mixture of tagged reference nucleic acids for propagation, comprising:
synthesizing multiple single-stranded reference nucleic acids;

tagging each of the multiple reference nucleic acids to make tagged reference nucleic acids, comprising:

5 synthesizing a copy of a first nucleic acid tag bound to a first end of each of the multiple reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer set and

10 synthesizing a copy of a second nucleic acid tag bound to a second end of each of the multiple reference nucleic acids, wherein each copy of the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

42. The method as recited in claim 41, wherein each of the multiple reference nucleic acids in the mixture of reference nucleic acids represent a different mutation of a normal human nucleic acid.

15

43. The method as recited in claim 41, wherein each of multiple reference nucleic acids in the mixture of reference nucleic acids includes at least one base that represents a mutation of a cystic fibrosis transmembrane conductance regulator gene for cystic fibrosis mutation testing.

20

44. The method as recited in claim 43, wherein each of the multiple reference nucleic acids in the mixture represents a different mutation of the cystic fibrosis transmembrane conductance regulator gene.

45. The method as recited in claim 41, wherein each reference nucleic acid in the mixture includes approximately the same number of bases.

5 46. The method as recited in claim 41, wherein for each of the multiple reference nucleic acids, approximately the same number of bases is used between a part of each reference nucleic acid representing a mutation and a part of each reference nucleic acid representing a nucleic acid tag.

10 47. The method as recited in claim 46, wherein at least one of the multiple reference nucleic acids in the mixture includes bases that represent multiple mutations of a normal human nucleic acid.

 48. The method as recited in claim 47, wherein approximately the
15 same number of bases are used between a part of a reference nucleic acid representing a first mutation and a part of the reference nucleic acid representing a second mutation as are used between each part of the reference nucleic acid representing a mutation and a part of each reference nucleic acid representing a nearest nucleic acid tag.

20

 49. The method as recited in claim 47, wherein if the mixture includes multiple reference nucleic acids each representing multiple mutations, then the number of bases separating bases representing different mutations is approximately equal for each of the multiple reference nucleic
25 acids representing multiple mutations.

50. The method as recited in claim 41, wherein a reference nucleic acid representing multiple mutations of a nucleic acid is ligated from at least two reference nucleic acids each representing a single mutation of the gene.

5 51. A tagged reference nucleic acid for a polymerase chain reaction amplification, comprising:

a synthesized reference nucleic acid having a base sequence capable of being used as a reference;

10 a first nucleic acid tag bound to a first end of the synthesized reference nucleic acid, wherein the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer set; and

a second nucleic acid tag bound to a second end of the synthesized reference nucleic acid, wherein the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

15

52. The tagged reference nucleic acid as recited in claim 51, wherein the synthesized reference nucleic acid includes a base sequence representing a mutation of a gene.

20 53. The tagged reference nucleic acid as recited in claim 52, wherein the gene comprises a cystic fibrosis transmembrane conductance regulator gene.

25 54. A mixture of tagged reference nucleic acids for a polymerase chain reaction amplification, comprising:

multiple synthetic reference nucleic acids;

a copy of a first nucleic acid tag bound to a first end of each of the multiple synthetic reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer set; and

5 a copy of a second nucleic acid tag bound to a second end of each of the multiple synthetic reference nucleic acids, wherein each copy of the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

10 55. A PCR amplifiable mixture of synthetic reference nucleic acids for genetic testing, comprising:

multiple synthetic reference nucleic acids each including a base sequence that represents a different site on a gene;

a copy of a first nucleic acid tag bound to a first end of each of the
15 multiple synthetic reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer set; and

a copy of a second nucleic acid tag bound to a second end of each of the multiple synthetic reference nucleic acids, wherein each copy of the
20 second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

56. One or more pairs of base sequences for producing one or more pairs of oligomers to act as nucleic acid priming sites for propagating
25 one or more mixtures of synthetic reference nucleic acids capable of being combined in a single mixture and amplified in a polymerase chain reaction

using one or more primer sets corresponding to the one or more base sequences, comprising:

5 a unique first base sequence for each of the one or more pairs of base sequences to produce copies of a first type of oligomer for use as a first type of nucleic acid priming site, wherein an end of each of the first type of oligomer is bound to a first end of each of the reference nucleic acids in the mixture of reference nucleic acids and wherein the first base sequence is complementary to a base sequence of a first primer of the two primers; and

10 a unique second base sequence for each of the one or more pairs of base sequences to produce copies of a second type of oligomer for use as a second type of nucleic acid priming site, wherein an end of each of the second type of oligomer is bound to a second end of each of the reference nucleic acids in the mixture of reference nucleic acids and wherein the second base sequence matches a base sequence of a second primer of the two
15 primers.

57. The pair of base sequences as recited in claim 56, wherein if one or more of the synthetic reference nucleic acids in the mixture represent one or more mutations, then each base sequence of each member of the pair
20 is distinct from base sequences of the synthetic reference nucleic acids in the mixture and distinct from each base sequence representing a mutation.

58. The pair of base sequences as recited in claim 56, wherein each of the base sequences in the pair are derived from a non-human source
25 for design of tags to be bonded to a synthetic reference nucleic acid associated with a human gene.

59. A primer set for amplifying a mixture of synthetic reference nucleic acids in a polymerase chain reaction, comprising:

5 a first primer having a base sequence complementary to a nucleic acid priming site bound to a first end of each of the synthetic reference nucleic acids in the mixture of reference nucleic acids; and

a second primer having a base sequence matching a nucleic acid priming site bound to a second end of each of the synthetic reference nucleic acids in the mixture of reference nucleic acids.

10

60. A kit including one or more primer pair designs and corresponding tag designs for manufacturing multiple mixtures of synthetic reference nucleic acids within a single mixture capable of being amplified by a polymerase chain reaction, comprising:

15 for each of the multiple mixtures: first and second tag designs for nucleic acid oligomers each having a base sequence unique over base sequences included in the synthetic reference nucleic acids in the mixture;

a first primer design:

20 wherein the first primer has a base sequence complementary to the base sequence of the first tag,

wherein copies of the first tag are bondable to a first end of each synthetic reference nucleic acid in the mixture, and

wherein each copy of the first tag comprises a nucleic acid priming site for the first primer; and

25 a second primer design:

wherein the second primer has a base sequence matching the base sequence of the second tag,

wherein copies of the second tag are bondable to a second end of each synthetic reference nucleic acid in the mixture, and

5 wherein each copy of the second tag provides a template for the first primer to produce a nucleic acid priming site for the second primer.

61. A method, comprising:

10 synthesizing multiple reference nucleic acids, wherein each of the multiple reference nucleic acids includes at least one mutation in a base sequence of a cystic fibrosis transmembrane conductance regulator (CFTR) gene;

15 synthesizing a copy of a first nucleic acid tag bound to a first end of each of the multiple reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer pair;

20 synthesizing a copy of a second nucleic acid tag bound to a second end of each of the multiple reference nucleic acids, wherein each copy of the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

62. The method as recited in claim 61, further comprising inserting approximately a consistent number of bases between each base representing a mutation on one of the reference nucleic acids and one of the tags on the reference nucleic acid.

25

63. The method as recited in claim 62, further comprising subjecting the mixture of tagged reference nucleic acids to a polymerase chain reaction in the presence of the first primer and the second primer.

5 64. A mixture of reference nucleic acids, comprising:
multiple reference nucleic acids, wherein each of the multiple reference nucleic acids is capable of being used as a clinical reference;

a copy of a first nucleic acid tag bound to a first end of each of the multiple reference nucleic acids, wherein each copy of the first nucleic acid tag has a base sequence complementary to a base sequence of a first primer of a primer pair; and

10

a copy of a second nucleic acid tag bound to a second end of each of the multiple reference nucleic acids, wherein each copy of the second nucleic acid tag has a base sequence matching a base sequence of a second primer of the primer set.

15

65. The mixture of reference nucleic acids as recited in claim 64, further comprising total genomic DNA from a reference human source that is known to differ at least one locus from one or more of the multiple reference nucleic acids.

20

66. The mixture of reference nucleic acids as recited in claim 65, wherein the total genomic DNA from a reference human source includes no known mutations.

67. A system for genetic mutation testing, comprising:
a means for inputting a patient genetic sample;
a reference nucleic acid mixture;
a means for comparing the patient genetic sample against at least one
5 constituent of the reference nucleic acid mixture; and
a means for displaying the comparison.

68. The system for genetic mutation testing as recited in claim 67,
wherein the reference nucleic acid mixture is used to calibrate at least one
10 part of the system.

69. The system for genetic mutation testing as recited in claim 67,
wherein the reference nucleic acid mixture is used to verify at least one part
of the system.

15

70. The system for genetic mutation testing as recited in claim 67,
wherein the reference nucleic acid mixture is used to verify another reagent
used in the system.

20

71. A method, comprising:
synthesizing a first mixture of various reference nucleic acids,
wherein each of the various reference nucleic acids in the first mixture
includes one or more tags allowing PCR amplification of the first mixture
via a primer set specific to the tags of the first mixture; and

25

synthesizing a second mixture of various reference nucleic acids,
wherein each of the various reference nucleic acids in the second mixture

includes one or more tags allowing PCR amplification of the second mixture via a second primer set specific to the tags of the second mixture.

72. The method as recited in claim 71, further comprising
5 combining the first and second mixtures to make a single mixture and differentially amplifying the first mixture and the second mixture in a PCR reaction by controlling amounts of the first primer set and the second primer set in the single mixture.

10 73. The method as recited in claim 72, wherein at least some of the reference nucleic acids include mutations of a normal human nucleic acid.

74. The method as recited in claim 73, further comprising adding
normal human nucleic acid to the single mixture to obtain heterozygous
15 pairs, wherein each heterozygous pair includes a normal segment of human nucleic acid and a mutated copy of the normal segment of human nucleic acid.

ABSTRACT

Subject matter includes design, synthesis, and propagation of synthetic and semi-synthetic reference nucleic acids and mixtures of reference nucleic acids for use in genetic tests, such as molecular screening, mutation testing, carrier testing, and diagnostic assays. In one implementation, methods are described for design, synthesis, and propagation of reference nucleic acid mixtures and a system is presented for using the mixtures.

Appendix A:

Synthetic Control Design Parameters: Exemplary PCR Components and Reaction Parameters

1. Components (25µl reaction)

dH ₂ O	14.5µl
10X PCR Buffer	2.5µl
MgCl ₂ (25mM)	2.5µl
dNTPs (2.0 mM)	2.5µl
Primers (10µM)	1.0µl
Amplitaq Gold 0.6U/µl)	1.0µl
Template (10µM mix)	1.0µl

2. Reaction Parameters

94°	10 minutes	
94°	45 seconds	5 Cycles
60°	45 seconds	
72°	45 seconds	
94°	45 seconds	25 Cycles
60°	45 seconds	
72°	45 seconds	
72°	5 minutes	
4°	∞	

Appendix B:

Synthetic Control Design Parameters: Exemplary Ligation Extension Components and Reaction Parameters

1. Components (20 μ l reaction)

CF ligation mixture	10 μ l
Bridge template (PCR amplicon / concentrated bridge oligo)	5.0 μ l
Reference nucleic acid mix (0.5 μ M)	5.0 μ l

2. Reaction Parameters

90°	5 seconds	32 Cycles
46.5°	45 seconds	
99°	10 minutes	
4°	∞	

Appendix C:

Synthetic Control Design Parameters: Exemplary Overlap Extension Components and Reaction Parameters

1. Components (25 μ l reaction)

dH ₂ O	16.5 μ l
10X PCR Buffer	2.5 μ l
MgCl ₂ (25mM)	2.5 μ l
dNTPs (2.0 mM)	2.5 μ l
Oligonucleotides (10 μ M)	1.0 μ l
Amplitaq Gold 0.6U/ μ l)	1.0 μ l

2. Reaction Parameters

94°	10 minutes	
94°	45 seconds	5 Cycles
60°	45 seconds	
72°	45 seconds	
94°	45 seconds	25 Cycles
60°	45 seconds	
72°	45 seconds	
72°	5 minutes	
4°	∞	

SEQUENCE LISTING

<110> Bejjani, Bassem A
Christensen, Todd M

<120> Methods of Designing, Synthesizing, and Propagating Reference
Nucleic Acids

<130> SH1-0001US

<140> US 10/731,419

<141> 2003-12-09

<160> 18

<170> PatentIn version 3.3

<210> 1

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Arbitrary Example Sequence of Reference Nucleic Acid for Clinical
Reference

<400> 1

agctatttcgc tagccgaaat agcgg

25

<210> 2

<211> 16

<212> DNA

<213> Artificial Sequence

<220>

<223> Example Synthetic Sequence

<400> 2

ctggccgtcg ttttac

16

<210> 3

<211> 17

<212> DNA

<213> Artificial Sequence

<223> Example Synthetic Sequence

<400> 3

caggaaacag ctatgac

17

<210> 4

<211> 16

<212> DNA

<213> Example Universal M13 Forward Primer

<400> 4

gaccggcagc aaaatg

16

<210> 5

<211> 17

<212> DNA
 <213> Example Universal M13 Reverse Primer
 <400> 5
 caggaaacag ctatgac 17

<210> 6
 <211> 42
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Example Complementary Sequence
 <400> 6
 gtcctttgtc gatactgtcg ataagcgatc ggctttatcg cc 42

<210> 7
 <211> 41
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Example Strand from Duplex Synthesis
 <400> 7
 tcgataagcg atcggttcta tcgccgaccg gcagcaaaat g 41

<210> 8
 <211> 17
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Example Synthetic Tag Sequence
 <400> 8
 gtcctttgtc gatactg 17

<210> 9
 <211> 37
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Example Second Strand from Duplex Synthesis
 <400> 9
 caggaaacag ctatgacagc tattcgctag ccgaaat 37

<210> 10
 <211> 9
 <212> DNA
 <213> Artificial Sequence
 <220>
 <223> Example Synthetic First Tag

<400> 10
tgatgatga 9

<210> 11
<211> 8
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Synthetic Second Tag

<400> 11
ccggaatt 8

<210> 12
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Example First Oligonucleotide

<400> 12
agattcgcta gcc 13

<210> 13
<211> 13
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Second Oligonucleotide

<400> 13
gaaatcgtag cgg 13

<210> 14
<211> 12
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Complement Bridge Sequence

<400> 14
gatcggcttt ag 12

<210> 15
<211> 26
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Ligation Sequence

<400> 15
agattcgcta gccgaaatcg tagcgg 26

<210> 16
<211> 6
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Synthetic Overlap Extension Complementary Strand End

<400> 16
cttttag

6

<210> 17
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Strand Sequence from Overlap Extension Method

<400> 17
gtcctttgtc gatactgaga ttcgctagcc ctttagcatc gccgaccggc agcaaaatg 59

<210> 18
<211> 59
<212> DNA
<213> Artificial Sequence

<220>
<223> Example Second Strand Sequence from Overlap Extension Method

<400> 18
caggaaacag ctatgactct aagcgatcgg gaaatcgtag cggctggccg tcgttttac 59